INVERSE SCATTERING AT LOW AND INTERMEDIATE FREQUENCIES

K.W. Fertig and J.M. Richardson
Rockwell International Science Center
Thousand Oaks, California 91360

ABSTRACT

In this paper, we address the inversion problem in the scattering of elastic waves from scatterers using a probabilistic approach. This work extends that of previous efforts in that it considers a wider set of input measurements and a wider class of possible defects. For a given transducer arrangement, the input measurements involve the coefficient A_2, which characterizes scattering at low frequencies, and a second quantity that is related to the distance from the center of the scatterer to the front face tangent plane perpendicular to the direction of the incident beam. The first property is deducible from low frequency scattering data and the second from low and intermediate frequency scattering data. These properties are determined for a set of transducer configurations. The class of possible scatterers now includes a finite discrete set of possible inclusions as well as a void. The boundary geometry is assumed to be ellipsoidal. In the probabilistic approach we start with a statistical ensemble of scatterer properties and measurement errors and then remove the members inconsistent with the scattering data obtained from the measurements. The best estimates of the geometrical properties and inclusion types (with the void regarded as a special case) are then the average or most probable values of these properties in the resultant reduced ensemble. These estimates are accompanied by several types of confidence measures. The behavior of the inversion algorithm using theoretical test data, both noiseless and noisy, was studied by computer simulation.

INTRODUCTION

In a previous report[1] we discussed the inverse problem in the scattering of elastic waves from voids at low frequencies (i.e., the Rayleigh regime). Here we consider an extended version of the inverse problem in which a wider set of types of measurements and types of scatterers are assumed. Namely, we consider a set of measurements that includes the low frequency scattering amplitudes as before but also the distance from the geometrical centers of the scatterer to the front-face tangent plane, a property deducible from the low and intermediate frequency characteristics of the scattering amplitude. As before a diversity of longitudinal-to-longitudinal scattering measurements are considered but this time attention is limited to the pulse-echo type. The set of possible scatterers has been extended to include a specified discrete set of possible inclusions as well as a void. As before, the boundary geometry is assumed to be ellipsoidal, although the actual calculations were performed for the oblate spheroidal case.

As before we pursue a probabilistic approach in which we start with a statistical ensemble of scatterer properties and measurement errors and then remove the members inconsistent with the scattering data obtained from the measurements. The best estimates of the geometrical properties and inclusion types (with the void regarded as a special case) are then the average or most probable values of these properties in the resultant reduced ensemble.

FORMULATION OF THE PROBLEM

In this section we present an outline of the analysis of the present extended version of the inverse scattering problem. In the next section, a discussion is given of the numerical approach and certain relevant analytical details. In the section following that, a set of preliminary results is presented giving the response of the inversion algorithm to various kinds of synthetic test data.

The possible results of scattering measurements (including post-detection processing) are represented by the stochastic mathematical model

$$y = h_g(x) + \nu \,, \tag{2.1}$$

where y and ν are N-dimensional vectors whose components are the possible measured values and measurement errors, respectively. The vector function $h_g(x)$ gives the error-free values of measurements that would be obtained with an ellipsoidal scatterer with inclusion type g ($g = 1, \cdots, G$) and geometry defined by the m-dimensional state vector x.

We consider two types of measurements and thus it is appropriate to write

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \tag{2.2}$$

$$\nu = \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \tag{2.3}$$

$$h_g = \begin{pmatrix} h_{g_1} \\ h_{g_2} \end{pmatrix} \,, \tag{2.4}$$

where y_1, ν_1, and h_{g_1}, are N_1-dimensional vectors associated with measurements of Type 1 and, similarly, y_2, ν_2, and h_{g_2}, are N_2-dimensional vectors associated with measurements of Type 2. Clearly, we must require $N_1 + N_2 = N$. In the case of Type 1 measurements, we obtain

$$h'_{g_1} = \left(h_{g_{11}}, \cdots, h_{g_{1N_1}} \right) \,, \tag{2.5}$$

where the prime denotes the transpose. The n^{th} component is defined by

$$h_{g_{1_n}} = h_{g_{1_n}}(x) = A_2(\vec{e}_n^s, \vec{e}_n^i; x, g) \ , \qquad (2.6)$$

where A_2 is the coefficient of ω^2 in the ω-expansion of the theoretical amplitude $A(\vec{e}_n^s, \vec{e}_n^i; x, g)$ for longitudinal-to-longitudinal scattering from an ellipsoidal inclusion of Type g and geometry x. The unit vectors $\vec{e}_n^i$ and $\vec{e}_n^s$ give the incident and scattered directions corresponding to the transducer placement in the n^{th} experiment (in the pulse-echo case $\vec{e}_n^s = -\vec{e}_n^i$).

In the Type 2 measurements, we have

$$h'_{g_2} = \left(h_{g_{21}}, \cdots, h_{g_{2N_2}}\right) \qquad (2.7)$$

in which the n^{th} component is defined by

$$h_{g_{2_n}} = h_{g_{2_n}}(x) = d(\vec{e}_n; x) \ , \qquad (2.8)$$

where d is the distance from the geometrical center of the ellipsoidal to a front-face tangent plane perpendicular to $\vec{e}_n$. The experimental value of this quantity is determined from a suitable analysis of the n^{th} pulse-echo scattering measurement in which $\vec{e}_n^s = -\vec{e}_n^i = \vec{e}_n$. It is worthy of note that in this formulation d depends only on x and not on g. A simple extension of this theory allows for dependence of d on g.

With x and g given, the experimental error vector ν is assumed to be a Gaussian random vector with the properties

$$E(\nu|x,g) = 0 \qquad (2.9)$$

$$E(\nu\nu'|x,g) = C \ . \qquad (2.10)$$

The $N \times N$ covariance matrix C is assumed to be independent of x and g. In the actual computations we assumed that C is diagonal, an assumption corresponding to the statistical independence of experimental errors.

The above stochastic model defines the probability density of y given x and g, namely $P(y|x,g)$ given by Eq. (3.1) of the next section. The statistical description of the measurement model is completed by the specification of the prior probability $P(x,g)$ which is given a more detailed discussion in the next section. The final decision (estimation of defect type and geometry) depends upon the posterior probability density of x and g given y, namely

$$P(x,g|y) = P(y|x,g)P(x,g)/P(y) \ , \qquad (2.11)$$

where

$$P(y) = \sum_g \int dx \ P(y|x,g)P(x,g) \qquad (2.12)$$

plays the role of a normalization constant.

The process of making a best decision about the values of x and g from the measurement vector y depends on the global context in which this problem is embedded. At this point it is necessary to focus the reader's attention on both the immediate as well as the ultimate purpose of the present inversion problem. The ultimate purpose involves the embedding of the inversion problem in a more extensive decision logic terminating in accept/reject decisions, estimation of life-cycle costs, etc. For the time being, however, it is expedient to regard the inversion algorithm as a "free-standing" entity with its own performance criteria.

With this point of view it is appropriate to consider a loss function $L(\hat{x},\hat{g};x,g)$ specifying the loss (or penalty) incurred if the algorithm gives the decisions $\hat{x}$ and $\hat{g}$ when the actual values are x and g. A reasonable optimization criterion is to minimize the risk defined by

$$R = EL(\hat{x}(y),\hat{g}(y);x,y) \ , \qquad (2.13)$$

where E is the unconditional averaging operator. It should be noted that here $\hat{x}$ and $\hat{g}$ are functions of the random process (2.1). Thus R is a functional of the decision functions $\hat{x}(y)$ and $\hat{g}(y)$ and it is to be minimized on the form of these functions.

Here we will consider two loss functions, namely

$$L(\hat{x},\hat{g};x,g) = -\delta(\hat{x} - x)\delta_{\hat{g}g} \qquad (2.14)$$

and

$$L(\hat{x},\hat{g};x,g) = -\delta_{\hat{g}g} \ . \qquad (2.15)$$

In the first case

$$E(L|y) = -P(x=\hat{x},g=\hat{g}|y) \qquad (2.16)$$

and in the second

$$E(L|y) = -P(g=\hat{g}|y) \ . \qquad (2.17)$$

Thus, in the first case the optimal decision functions correspond to the most probable values of x and g given y. The second case will be discussed later.

COMPUTATIONAL APPROACH

As discussed in the last section, the estimation of the inclusion type g and geometry x is performed using standard decision theory. Using either loss function, i.e., Eq. (2.14) or (2.15), we must consider the posterior probability of g and x given the observed measurement vector, y. Using the notation of the last section, we represent the measurement model (2.1) as the probability density

$$P(y|x,g) = \frac{1}{(2\pi)^{N/2}} |C|^{-1/2} \exp\left[-\frac{1}{2}\left(y - h_g(x)\right)'\right.$$
$$\left. \times C^{-1}\left(y - h_g(x)\right)\right] \ . \qquad (3.1)$$

The covariance matrix C was defined in the previous section and $|C|$ is its determinant.

The posterior probability of x and g is obtainable once the prior on x and g is specified. Denoting this prior by $P(x,g)$, the posterior probability is then

$$P(x,g|y) = P(y|x,g)P(x,g)/P(y) \ , \qquad (3.2)$$

where $P(y)$ plays the role of a normalizing constant. Maximization of (3.2) with respect to x and g is performed by sequentially maximizing (3.2) on x for each specified g, $g = 1, \cdots, G$. This latter maximization is conveniently performed using the International Mathematics and Statistics Library routine ZXMIN. This routine[2] uses a quasi-Newton algorithm to find the minimum of a function of m variables. In the current application, it was found that certain precautionary steps are necessary in order to assure convergence of the iterative scheme. Before discussing these steps, it is desirable to elaborate on the computational procedure to obtain $P(x,g|y)$ in (3.2) near its maximum.

Maximization of (3.2) over x for a given g is equivalent to minimizing the functional $\phi(x,g;y)$ over x for specified y where ϕ is defined by

$$\phi(x,g;y) = -\ell n[P(x,g|y)P(y)(2\pi)^{-N/2}|C|^{-1/2}]$$
$$= \frac{1}{2}\left(y - h_g(x)\right)'C^{-1}\left(y - h_g(x)\right) - \ell n\, P(x,y) \tag{3.3}$$

Specifically, ϕ is completely known without having to evaluate the normalization constant in (3.2). Let $\hat{x}_g(y)$ be the value of x for which the minimum is obtained. That is

$$\phi(\hat{x}_g,g;y) = \min_x \phi(x,g;y) \ . \tag{3.4}$$

Thus $\hat{x}_g = \hat{x}_g(y)$ gives the geometry of the inclusion that maximizes the posterior density function for inclusion type g. Since the gradient of ϕ with respect to x vanishes at $\hat{x}_g$, we see that $\phi(x,g;y)$ has the series expansion

$$\phi(x,g;y) = \phi(\hat{x}_g,g;y) + \frac{1}{2}(x - \hat{x}_g)'H_g(x - \hat{x}_g) + \cdots \ , \tag{3.5}$$

where

$$H_{g_{ij}} = \left[\frac{\partial^2\phi}{\partial x_i \partial x_j}\right]_{x=\hat{x}_g}$$

is the Hessian of the function ϕ at $x = \hat{x}_g$. A numerical approximation of H_g is computed in the inversion software using central difference approximations. We note that for ϕ to have a true minimum at $x = \hat{x}_g$, H_g must be positive definite.

The expansion in (3.5) allows for an approximate expression to be developed for $P(x,g|y)$ given in (3.2) using the relation (3.3). In particular, we obtain the expression

$$P(x,g|y) \simeq A(y) \exp\left[-\phi(\hat{x}_g,g;y)\ \tfrac{1}{2}(x-\hat{x}_g)'H_g(x-\hat{x}_g)\right], \tag{3.6}$$

where $A(y)$ is now easily computed as

$$A(y) = \left(\sum_g \int dx\ P(x,g|y)\right)^{-1}$$
$$= \left[(2\pi)^{m/2} \sum_{g=1}^{G} |H_g|^{-1/2} \exp\left(-\phi(\hat{x}_g,g;y)\right)\right]^{-1} \tag{3.7}$$

In Eq. (3.7), m is the number of elements in the geometric state vector x. The approximation in (3.6) is used to avoid the multidimensional numerical integration implied in (3.2) to define $P(y)$. The posterior probability of inclusion type can be found by integrating out x in (3.6). The result is

$$P(g|y) = \exp\left[-\phi(\hat{x}_g,g;y)\right]|H_g|^{-1/2}$$
$$\times\left[\sum_{g'=1}^{G} \phi(\hat{x}_{g'},g';y)|H_{g'}|^{-1/2}\right]^{-1} \tag{3.8}$$

The maximum value of $P(x,g|y)$ for a specified g is

$$P(\hat{x}_g,g|y) = A(y) \exp\left[-\phi(\hat{x}_g,g|y)\right] \ , \tag{3.9}$$

where the normalization factor $A(y)$ is given by Eq. (3.7). A measure of the uncertainty in x for any assumed g is given by the posterior covariance matrix of x for the value of g specified. Using the above approximations, we find this to be just $\mathrm{cov}(x|g,y) = H_g^{-1}$.

Two decision rules are of interest in choosing an inclusion type, g. The first is to choose the value of g that maximizes $P(g|y)$ in (3.8) corresponding to the loss function (2.15) in the last section. The second is to choose that value of g that maximizes $P(\hat{x}_g,g|y)$ in (3.9), which is consistent with the loss function (2.14) in the last section. The former rule has a certain intuitive appeal. If the present inversion procedure is regarded as a "free-standing" entity then either rule can be used, the corresponding loss functions being a matter of taste. However, if the procedure is to be integrated into a larger, more complex, scheme of decision logic yielding optimal accept/reject policies, etc., then no decisions are to be made in the inversion procedure, except as an approximation to simplify the processing in the remainder of decision logic. As an approximation, it is usually sufficient to input into the remaining decision logic the approximation to $P(x,g|y)$ given by (3.6).

At present, the computer program implementing the inversion algorithm reports both $P(g|y)$ and $P(\hat{x}_g,g|y)$ for each g as well as the posterior standard error of each component of x and correlation matrix of these components. The program is written in FORTRAN 5 and is able to run in pseudo real time on a Data General M-600. [Approximately two minutes of CPU time (timeshare mode) is required to estimate the geometry for each inclusion type in a class of five inclusions.] Currently, there are several options available for specification of the prior distribution of x and y. In each of these options, the geometry x is assumed to be distributed independently of the inclusion type, g. This restriction is easily relaxed with a small modification to the computer code. The discrete prior distribution of g is specified in the input. Normally, this will be taken as flat (equal probability over all inclusion types). The flow of the various types of input information is shown in Fig. 1 along with the major output discussed above.

In the case of spheroidal voids or inclusions, the prior distribution of $x = (a,c,\gamma_z,\alpha)$ is assumed factorable into the following three parts:

$$P(a,c,\gamma_z,\alpha) = P(a,c)P(\gamma_z)P(\alpha) \ . \tag{3.10}$$

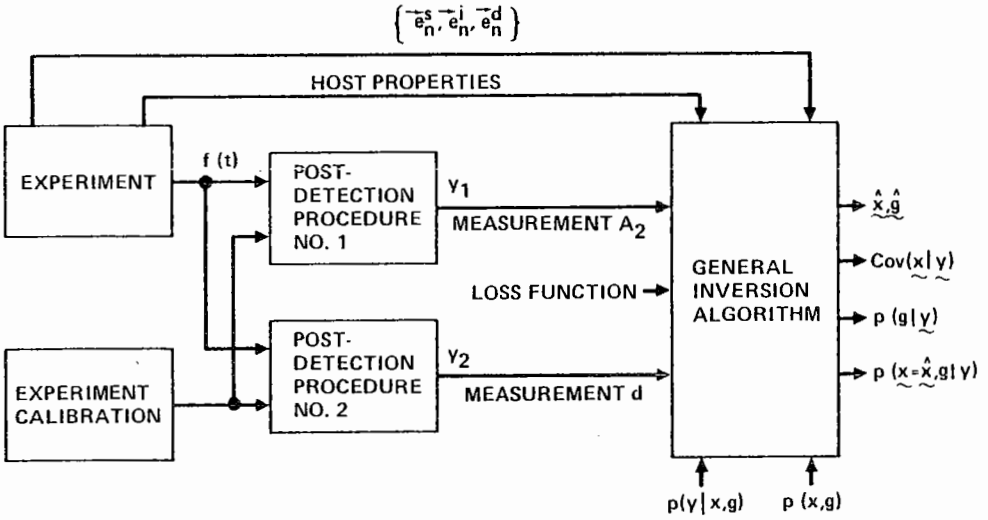

Fig. 1. Schematic of Information Flow for General Inversion Algorithm

In the above expression, a and c are the semi-major and semi-minor axis lengths of an oblate spheroid and γ_z and α define the orientation (i.e., the direction cosine relative to the z-axis and the azimuthal angle, respectively).

The program, as currently implemented, allows for the following parameterization of the prior distribution:

(1) $\gamma_z \sim$ Beta with input degrees of freedom v_{z_1} and v_{z_2}.

(2) $\alpha/2\pi \sim$ Beta with input degrees of freedom v_{α_1} and v_{α_2}.

(3) (i) (a,c) ~ Bivariate Weibull with input shape and scale parameters.

(ii) (a,c) ~ Bivariate lognormal with input location and scale parameters.

(iii) (a,c) ~ Bivariate extreme value type-I with input location and scale parameters.

(iv) (a,c) ~ Flat.

The joint distribution of a and c is restricted to have positive mass only when $a \geq c$. This is because the current coding of the low frequency scattering algorithm is restricted to oblate spheroidal geometry.

The beta priors on γ_z and x include the uniform density as a special case ($v_{z_1}=v_{z_2}=v_{\alpha_1}=v_{\alpha_2}=1$). Specifications of the uniform density for these parameters is equivalent to assuming *a priori* that the axis of symmetry is uniformly distributed over the unit sphere. If this uniform distribution is assumed and if the distribution of a and c is taken as flat (no *a priori* bias) the program will automatically produce the maximum likelihood estimate of x for each category g. This is because the loss function being used is the Dirac delta function. The posterior variance reported by the program becomes the conditional covariance matrix of x_g conditioned on x and g. The distributional properties of x_g in this conditional situation are discussed in Appendix C.

The inversion algorithm as currently implemented requires the measurement covariance matrix C as input. This matrix is assumed to be diagonal (independent observations). The program reads in an estimate of the standard error of y_{1_n} and y_{2_n} for each y_{1_n} and y_{2_n} input. This estimate will presumably come from replicate experiments or from various comparisons of theory and experiment. The actual estimated $\hat{x}_g$ depends only on the relative sizes of these standard errors, not their absolute magnitude. The posterior standard error does depend upon their magnitude however. The following is reported for each inclusion type considered:

1. Posterior probability of g: $P(g|y)$.
2. Posterior density of x and g at $\hat{x}_g$: $P(\hat{x}_g,g|y)$.
3. Residual sum of squares for A_2 measurements: $(y_1-h_1(\hat{x}_g))'C_1^{-1}(y_1-h_1(\hat{x}_g))$.
4. Residual sum of squares for d measurements: $(y_2-h_2(\hat{x}_g))'C_2^{-1}(y_2-h_2(\hat{x}_g))'$.
5. Total residual sum of squares: $(y-h_g(\hat{x}_g))'C^{-1}(y-h_g(\hat{x}_g))$.
6. Convergence information.
7. Estimates of the two major axes (a,a) and the one minor axis (c) and the direction cosines of all axes.
8. Posterior standard errors of all items in (7).
9. Posterior correlation matrix of all items in (7).
10. Observations, predicted values and residuals for each of the A_2 type measurements and each of the d type measurements.

All standard errors and correlations are computed using the usual linear approximation techniques. In particular, if v = h(u) where v and u are random vectors, then the corresponding covariance matrices are related by the approximate expression

$$E_v \simeq J\Sigma_u J' ,$$

where

$$J = \partial h/\partial u$$

is the Jacobian matrix of the transformation h. This matrix need not be square.

The residual sums of squares in outputs (3) through (5) above allow one to test the statistical model's adequacy. In particular, these sums of squares will be approximately χ^2 distributed when the inclusion type is correctly assumed. They give a direct measure of the experimental uncertainty. In Appendix C, we discuss their distributional form in more detail.

The residuals reported in output (10) provide a further indication of the model's adequacy. Simple run tests performed on these residuals will indicate non-random patterns and thus address the question of goodness-of-fit of the mathematical scattering model to the experimental situation for each postulated inclusion type. Those postulated inclusion types which exhibit non-random residuals should be considered as wrong candidates.

As mentioned earlier, certain precautionary steps are necessary in order to assure convergence of the quasi-Newton algorithm used by ZXMIN. First of all, it is useful to transpose the state vector $x = (a,c,\gamma_z,\alpha)$ into the state vector x^* given by

$$\begin{aligned} x_1^* &= \ell n\ (a/a^*) \\ x_2^* &= \ell n\ (c/(c^*a-c)) \\ x_3^* &= \gamma_x/\gamma_z \\ x_4^* &= \gamma_y/\gamma_z\ , \end{aligned} \tag{3.11}$$

where

$$\gamma_x = \cos\alpha \sin\theta$$

$$\gamma_y = \sin\alpha \sin\theta$$

and

$$\gamma_z = \cos\theta\ .$$

Here, a* is an arbitrary scaling constant chosen so that underflows or overflows of digits are avoided. The term c* is a biasing factor that is used to avoid singularities encountered in the calculation of A_2 when the geometry is too close to a sphere. It represents the maximum allowable ratio of c/a.

The transformation specified by (3.11) maps the original state variable x into an unconstrained state variable x*. In particular, each component of x* may range between $\pm\infty$. This mapping, which is 1-1, greatly facilitates the use of ZXMIN which would otherwise ignore the constraints in the original state space: $0 \leq \gamma_z \leq 1$, $0 \leq \alpha \leq 2\pi$, and $a \geq c \geq 0$.

The inverse transformation is given by

$$a_1 = a^* e^{x_1^*}$$

$$c = c^* a^* e^{x_1^*} e^{x_2^*}/(1+e^{x_2^*})$$

$$\gamma_x = x_3^*/(1+x_3^{*2}+x_4^{*2})^{1/2}$$

$$\gamma_y = x_4^*/(1+x_3^{*2}+x_4^{*2})^{1/2}$$

$$\gamma_z = 1/(1+x_3^{*2}+x_4^{*2})^{1/2}$$

In actual implementation of the inversion algorithm, it has been found that convergence is quite sensitive to the first guess. When there is little *a priori* bias, it often turns out that the first guess will be at a place on the response surface $\phi(x,g;y)$ where the Hessian is not positive definite. This implies that the surface at a small distance from the minimum has a ridge-canyon behavior that will defeat the algorithm if suitable tactics are not utilized. The iteration scheme used in ZXMIN employs an approximation to the Hessian. At the initiation of the algorithm, any positive definite matrix may be used (not necessarily the true Hessian) for this approximation. The approximation is updated each iteration and converges to the true Hessian as the true minimum is reached. The rapidity of convergence is tied to the difference between the initial approximate Hessian and the true one. If the true Hessian at x_0 is not positive then it cannot be used as a first guess. The program as implemented, forces this matrix to be positive definite by taking the absolute value of the diagonal portion of the Hessian when it is discovered that the Hessian is not computationally positive definite. In order to reduce computational effort as much as possible, as well as to assure that a local minimum is to be obtained, the program uses a simple, but very robust initial search algorithm to find the approximate location of a local minimum in the four dimensional space of x*. In the simulation runs performed to date, it was often found that the use of this robust search algorithm before employing ZXMIN was essential to assure convergence when the inclusion type assumed was different from that which was used in the generation of the test data. It is just as important that convergence be reached for incorrect inclusion types as for the correct ones, so that a proper assessment of posterior probabilities can be made.

NUMERICAL RESULTS

In this section, we describe some of the recent numerical testing of the inversion algorithm. The purpose is to determine the robustness of the algorithm to perturbations of the various assumptions in both the statistical model and the physical scattering model. As a minimum, it is necessary to assess the effect on the inversion results of 1) flaw sizes, 2) flaw orientation, 3) flaw type, 4) host property variation, 5) *a priori* distributional assumptions, 6) actual measurement error, 7) assumed rms measurement error, 8) transducer placements, and 9) scattering theory inaccuracies. To date, only some of the numerical experiments addressing items 1 through 7 have been run. It has been found that the inversion algorithm performs quite well over a wide range of experimental errors. Some of the results are reported in Table 1. This table is based on synthetic scattering data generated using a Gaussian random number generator. The standard errors reported as σ_{A_2} and σ_d under the subheading "Experimental" were used to scale the Gaussian random number. The inversion algorithm used the standard errors reported under the "Analysis" subheading. Three simulations were performed for each case. In point of fact, if the function ϕ being minimized

TABLE 1
INVERSION ALGORITHM NUMERICAL RESULTS FOR SELECTED AREAS (MAXIMUM LIKELIHOOD)

Experimental		Analysis								
σ_{A_2}	σ_d	σ_{A_2}	σ_d	a	σ_a	c	σ_c	γ_z	σ_{γ_z}	P(void)
	A. Void, a = 0.04, c = 0.02, γ_z = 1.0									
10%	20%	10%	20%	0.0405	0.0009	0.0203	0.0018	0.993	0.005	0.992
				0.0402	0.0009	0.0200	0.0019	0.982	0.015	0.989
				0.0400	0.0009	0.0205	0.0018	0.997	0.006	0.982
20%	40%	20%	40%	0.0412	0.0017	0.0203	0.0036	0.990	0.021	0.779
				0.0415	0.0020	0.0182	0.0038	0.917	0.064	0.790
				0.0400	0.0017	0.0209	0.0035	0.988	0.025	0.669
10%	20%	10%	10,000%	0.0410	0.0011	0.0196	0.0022	0.982	0.017	0.345
				0.0417	0.0010	0.0168	0.0021	0.977	0.016	0.307
				0.0396	0.0011	0.0212	0.0024	0.997	0.009	0.354
	B. Void, a = 0.04, c = 0.038, γ_z = 1									
10%	20%	10%	10%	0.0409	0.0009	0.0374	0.0018	0.793	0.335	0.995
				0.0422	0.0011	0.0344	0.0019	0.555	0.179	0.911
				0.0405	0.0009	0.0375	0.0018	0.739	0.407	0.997
5%	10%	10%	20%	0.0403	0.0008	0.0380	0.0018	0.937	0.243	0.998
				0.0406	0.0010	0.0370	0.0019	0.692	0.352	0.988
				0.0401	0.0008	0.0380	0.0017	0.921	0.306	0.999
	C. Void, a = 0.04, c = 0.038, γ_z = 0.5									
10%	20%	10%	20%	0.0412	0.0013	0.0367	0.0021	0.443	0.330	0.991
				0.0427	0.0012	0.0334	0.0019	0.428	0.152	0.826
				0.0413	0.0012	0.0359	0.0020	0.444	0.274	0.990
	D. Void, a = 0.04, c = 0.020, γ_z = 0.5, α = 45°									
10%	20%	10%	20%	0.0402	0.0012	0.0209	0.0017	0.489	0.079	0.840
				0.0409	0.0012	0.0189	0.0016	0.461	0.068	0.614
				0.0410	0.0012	0.0190	0.0015	0.507	0.068	0.939

to determine x_g were a pure quadratic form in x, then the posterior variances reported in Table 1 would not be a function of the specific random vector y observed, but only a function of the assumed values of σ_{A_2} and σ_d, the transducer placements, and the inclusion type. Deviation from quadratic behavior in the functional form of ϕ yields additional variability in the posterior variances.

All of the scattering measurements were taken to be of the pulse-echo type. The transducers were assumed to be placed in a symmetrical array on a spherical test piece with the inclusion at the center. Seventeen transducers were used. One was placed directly above the flaw. The remaining sixteen transducers were placed four each, 90° apart, in four different cones above the inclusion. The four cones had half angles of 15°, 30°, 45°, and 60°, respectively. The optimum number and placement of transducers is still under investigation.

All of the analyses represented in Table 1 assumed *a priori* that the flaw being measured was one of five types: void, BN, Si, SiC, or WC. Given the low frequency scatter coefficients, A_2, for SiC and WC are negative in a SiN_3 host, these two categories may be eliminated immediately when the true inclusion type is a void. The last column in Table 1 reports the posterior probability that the inclusion type is a void. The only serious contender using the five types allowed was BN. In every case run so far, the posterior probabilities of inclusion other than a void or BN (when a void was used in the test data) were all less than 0.0001. Similar results using a different random seed are plotted in Fig. 2 wherein the actual measurement error ranges from 5% to 20% in A_2 and from 10% to 40% in d.

The results reported in Part A of Table 1 and Fig. 2 indicate two things. First, very large random components of the measurement error may be tolerated before the inversion breaks down. The

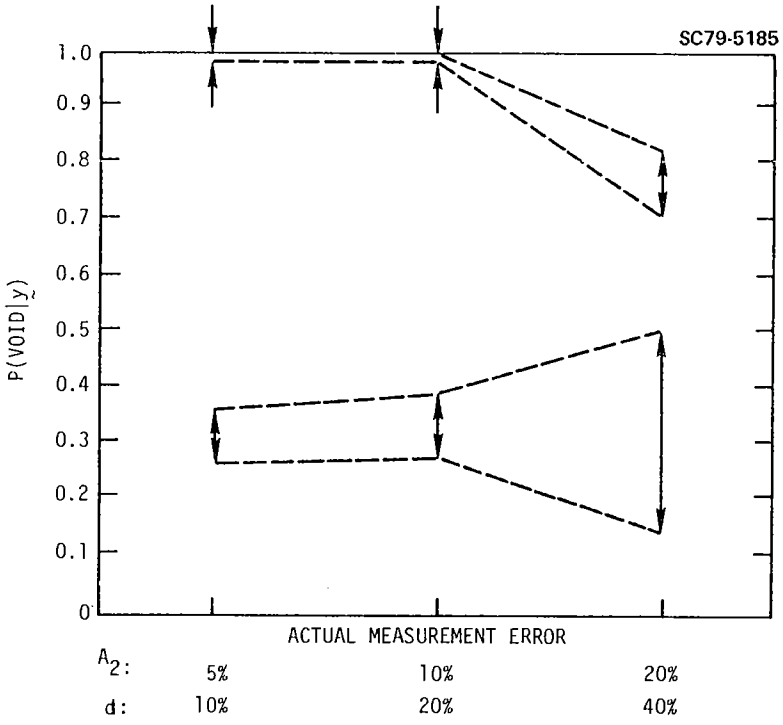

Fig. 2. Effect of Random Measurement Error On Classification Probability (Oblate Void 200 μm/400 μm)

posterior probability of 0.66 to 0.79, with 20% to 40% measurement error per measurement, is quite good. (Note that since 17 observations are being taken for each of the A_2 type measurements, the effective random error in inversion is approximately $17^{-1/2}$ or 0.24 of the individual errors.) The Gaussian approximation to the posterior distributions of a and g for the various flaw types is shown in Fig. 3 for a specific set of random measurement errors. The area under each flaw's curve represents the posterior probability of that flaw type. The spread of the curve indicates the posterior uncertainty in the major semi-axis length. It can be seen that the void is by far the most probable flaw category in the three cases shown. Also, the spread widens as the measurement increases. This is a natural result.

The second item of interest seen from Part A and Fig. 2 is that it appears that the d type measurements are crucial to the inversion procedure in most cases. The run whose σ_d was taken as 10,000% in the analysis stage effectively gives the d measurements zero weight for that inversion. The result is an inversion based on A_2 measurements alone. As can be seen, the posterior probability of a void is only 0.3 to 0.36. In fact, the posterior probability of BN was 0.60 to 0.70 for these runs. On the other hand, it is apparent from the estimate of a, c, and γ_z, that the geometry is still being estimated quite accurately based on the A_2 measurements alone. The unfortunate part is that the BN category seems to fit the noisy A_2 data for a void just as well as a void does. In the actual three simulation runs, the major difference between the BN and void computation was in $|H_g|^{-1/2}$. This is a measure of the curvature of the posterior density of x near its maximum. This curvature was larger for the BN category than for the void category. Interestingly, both categories produced essentially the same residual sum of squares.

Part B of Table 1 represents some of the results obtained when the inclusion becomes nearly spherical. As can be seen, the posterior probability of a void remains high. The posterior standard error, σ_{γ_z}, becomes quite large. This is to be expected, since in the limit when the flaw becomes a sphere, the axis of symmetry loses meaning.

Part C of the table demonstrates that tilting a nearly spherical flaw has essentially no effect on the inversion algorithm. However, tilting a more eccentric spheroid does, of course, have an effect. Part D of the table represents the results of the inversion algorithm in the case where the axis of symmetry of a flaw with dimensions a = 0.04 cm and c = 0.02 cm is 45° off of the vertical with an azimuthal angle of 45°. Further results using a different random seed are presented in Fig. 4. For certain orientations, the classification probability becomes quite poor. Presumably, this difficulty is correlated with the transducer configuration used. Clearly, it is desirable to devote more work to the quantification of the effect of transducer placement.

No runs have been made yet to assess the effect of scattering model error. In particular, it is desirable to determine the effect on inversion of such things as, 1) losses in the host, 2) multiple scatterers, and 3) non-spheroidal scatterer geometry.

REFERENCES

1. R.K. Elsley, J.M. Richardson and R.B. Thompson, "Determination of Fracture Mechanism Parameters from Elastic Wave Scattering at Lower Frequencies," Semi-Annual Report, Interdisciplinary Program for Quantitative Flaw Definition, Report SC595.32SA, pp. 104-14 , February 1978.
2. Fletcher, R., "Fortran Subroutine for Minimizatior by Quasi-Newton Methods," R7125 AFRF, Harwell, UK.

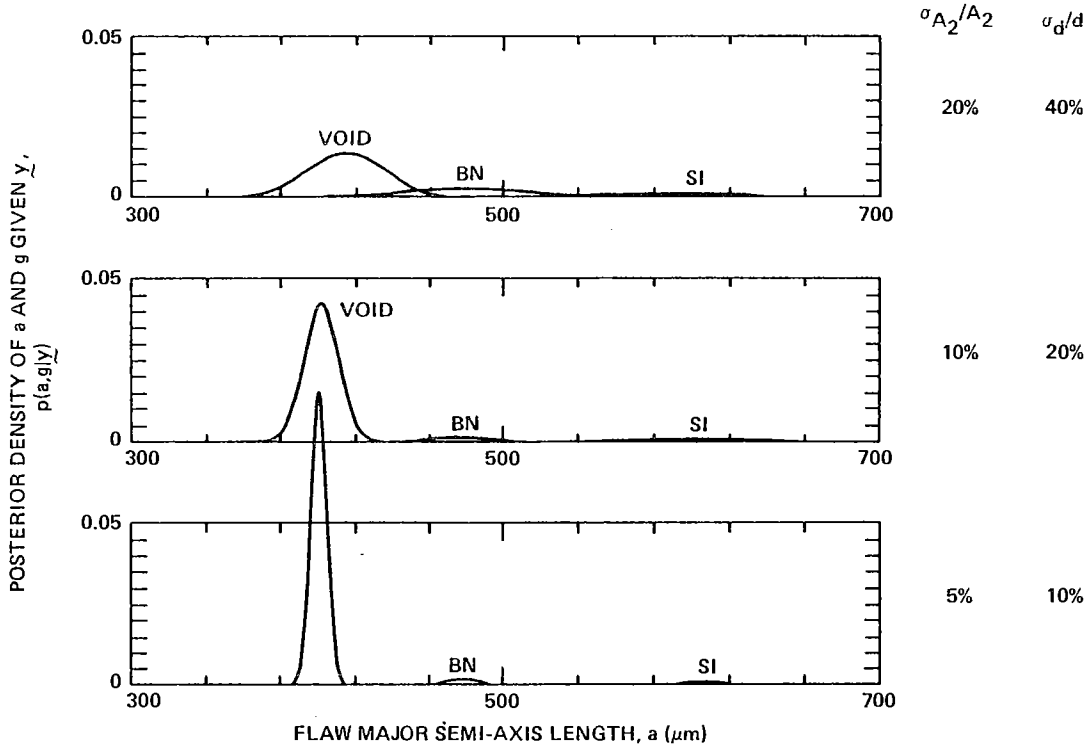

Fig. 3. Effect of Measurement Error on Estimated Geometry of Ellipsoidal Inclusions (Oblate Void 200 μm/400 μm)

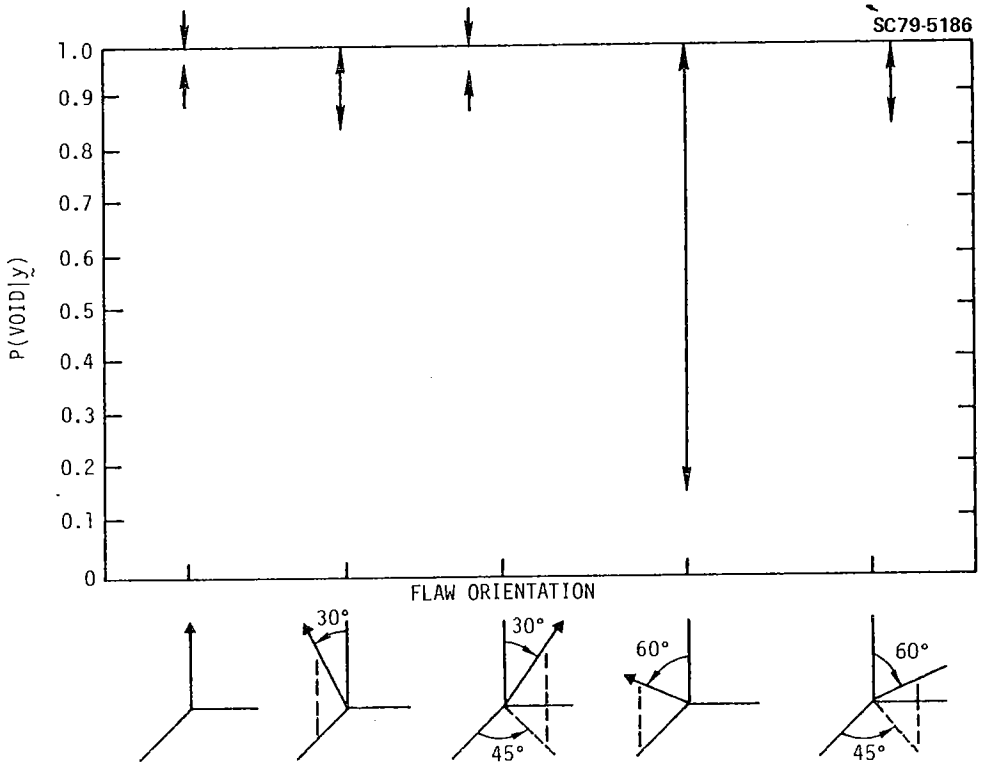

Fig. 4. Effect of Orientation on Classification Probability (Oblate Void 200 μm/400 μm)

APPENDIX A

SCATTERING OF ELASTIC WAVES FROM ELLIPSOIDAL VOIDS AND INCLUSIONS IN THE LOW FREQUENCY (RAYLEIGH) REGIME

It is our purpose here to present further details of the analysis of $A_2 = A_2(\vec{e}^s,\vec{e}^i;x,g)$ for the case of longitudinal-to-longitudinal elastic wave scattering. In this case, we have that

$$A^{(2)}_{\ell-\ell}(\vec{e}^s,\vec{e}^i) = \frac{V_s}{4\pi\rho}\left[\delta\rho c_\ell^{-2} e^s_\alpha e^i_\alpha - c_\ell^{-4} t^s_m (G_{mn} + \delta C^{-1}_{mn})^{-1} t^i_n\right] , \qquad \text{(A-1)}$$

where repeated subscripts imply summation. In the following, the subscripts α,β,γ, and δ will range from one to three, representing the x,y, and z directions in a laboratory coordinate system, while the subscripts m and n range from one to six.

In equation (A-1), V_s is the volume of the ellipsoid and is given by

$$V_s = \frac{4\pi}{3} abc ,$$

where a, b and c are the semiaxes lengths of the ellipsoid with principal axes defined by the unit vectors $\vec{u}$, $\vec{v}$, and $\vec{w}$, respectively. The parameter c_ℓ is the velocity of longitudinal waves in the bulk and is given by

$$c_\ell = [(\lambda + 2\mu)/\rho]^{\frac{1}{2}} ,$$

where ρ is the density of the bulk and λ and μ are the Lamé constants of the bulk.

We let $\rho + \delta\rho$, $\lambda + \delta\lambda$, $\mu + \delta\mu$ be the density and Lamé constants of the inclusion.

The six vectors t^s_m and t^i_m are obtained via a transformation from the 3×3 matrix representation of strain. Specifically,

$$t^s_m = e^s_\alpha e^s_\beta s_{m\alpha\beta} ,$$

$$t^i_m = e^i_\alpha e^i_\beta s_{m\alpha\beta} ,$$

where $\vec{e}^s = (e^s_\alpha)$ is the unit vector in the scattered direction and $\vec{e}^i = (e^i_\alpha)$ is the unit vector in the incident direction of the plane acoustic wave.

Recalling that $\vec{u} = (u_\alpha)$, $\vec{v} = (v_\alpha)$, $\vec{w} = (w_\alpha)$ are the three principal axes of the ellipsoid, we have that the transformation matrix $\vec{s}$ is given by

$$s_{1\alpha\beta} = u_\alpha u_\beta$$

$$s_{2\alpha\beta} = v_\alpha v_\beta$$

$$s_{3\alpha\beta} = w_\alpha w_\beta$$

$$s_{4\alpha\beta} = \frac{1}{\sqrt{2}}(v_\alpha w_\beta + w_\alpha v_\beta)$$

$$s_{5\alpha\beta} = \frac{1}{\sqrt{2}}(w_\alpha u_\beta + u_\alpha w_\beta)$$

$$s_{6\alpha\beta} = \frac{1}{\sqrt{2}}(u_\alpha v_\beta + v_\alpha u_\beta) ,$$

where $\alpha,\beta = 1,\cdots,3$.

The six by six matrix δC_{mn} is given by

$$\delta C_{mn} = s_{m\alpha\beta}\delta C_{\alpha\beta\gamma\delta} s_{n\gamma\delta} ,$$

where $\delta C_{\alpha\beta\gamma\delta}$ is the elastic constant deviation tensor in the inclusion.

The Green's matrix G_{mn} has very special structure for an oblate spheroid ($a = b > c$). It is given by

$$G_{mn} = \left(\begin{array}{ccc|ccc} G_{11} & G_{12} & G_{13} & & & \\ G_{12} & G_{11} & G_{13} & & 0 & \\ G_{13} & G_{13} & G_{33} & & & \\ \hline & & & G_{44} & 0 & 0 \\ & 0 & & 0 & G_{44} & 0 \\ & & & 0 & 0 & G_{66} \end{array}\right) ,$$

with

$$G_{11} = pI_a - q(2I_a - 3a^2 I_{aa})$$

$$G_{33} = p(4\pi - 2I_a) - q(8\pi - 4I_a - 3c^2 I_{cc})$$

$$G_{12} = -q(I_a - a^2 I_{aa})$$

$$G_{13} = -q(I_a - 3c^2 I_{ac})$$

$$G_{44} = \frac{1}{2}p(4\pi - I_a) - \frac{1}{2}q(4\pi - I_a - 6(a^2+c^2)I_{ac})$$

and

$$G_{66} = pI_a - q(I_a - 2a^2 I_{aa}) .$$

In the above,

$$I_a = \frac{2\pi a^2 c}{(a^2-c^2)^{3/2}}\left(\cos^{-1}\frac{c}{a} - \frac{c}{a}\left(1 - \frac{c^2}{a^2}\right)^{\frac{1}{2}}\right)$$

$$I_{aa} = \frac{3}{4(a^2-c^2)} I_a - \frac{\pi c^2}{a^2(a^2-c^2)}$$

$$I_{ac} = \frac{4\pi - 3I_a}{3(a^2-c^2)}$$

and

$$I_{cc} = \frac{2}{a^2-c^2} I_a + 4\pi \frac{a^2 - 3c^2}{3c^2(a^2-c^2)} ,$$

with

$$p = \frac{1}{8\pi\rho} \left(\frac{1}{c_t^2} + \frac{1}{c_\ell^2}\right)$$

and

$$q = \frac{1}{8\pi\rho} \left(\frac{1}{c_t^2} - \frac{1}{c_\ell^2}\right) .$$

Here c_t is the transverse wave velocity in the host and is given by

$$c_t = (\mu/\rho)^{\frac{1}{2}} .$$

In the special case of a sphere (a=b=c), the above is somewhat simpler. Specifically,

$$G_{11} = G_{33} = \frac{3 + \kappa}{15\rho c_\ell^2}$$

$$G_{12} = G_{13} = \frac{2 - \kappa}{30\rho c_\ell^2}$$

$$G_{44} = G_{66} = \frac{4 + 3\kappa}{30\rho c_\ell^2}$$

with

$$\kappa = 2(\lambda + 2\mu)/\mu .$$

APPENDIX B

DISTANCE FROM THE CENTER TO THE FRONT-FACE TANGENT PLANE FOR ELLIPSOIDAL GEOMETRY

We present here a detailed derivation of the distance $d(\vec{e};x)$ discussed in the second section of the paper. The general equation of an ellipsoid can be written in the form

$$\phi(\vec{r}) \equiv \vec{r} \cdot \overleftrightarrow{Q} \cdot \vec{r} = 1 \ , \tag{B-1}$$

where the tensor $\overleftrightarrow{Q}$ is given by

$$\overleftrightarrow{Q} = a^{-2}\vec{u}\vec{u} + b^{-2}\vec{v}\vec{v} + c^{-2}\vec{w}\vec{w} \tag{B-2}$$

in which a, b and c are the semi-axis lengths and $\vec{u}$, $\vec{v}$ and $\vec{w}$ are the mutually orthogonal unit vectors giving the directions of the corresponding principal axes. In the case of a spheroid we obtain

$$\begin{aligned}\overleftrightarrow{Q} &= a^{-2}(\vec{u}\vec{u} + \vec{v}\vec{v}) + c^{-2}\vec{w}\vec{w} \\ &= a^{-2}(\overleftrightarrow{1} - \vec{w}\vec{w}) + c^{-2}\vec{w}\vec{w} \ ,\end{aligned} \tag{B-3}$$

where $\overleftrightarrow{1}$ is the unit tensor. It is clear that $\phi(\vec{r})$ depends upon the geometrical state vector x, but only through the tensor $\overleftrightarrow{Q}$.

Let us consider a plane that is tangent to the ellipsoid at the point $\vec{r}$ and that has an outward pointing normal $\vec{e}$. This plane represents a wavefront impinging on the ellipsoid in a pulse-echo experiment in which $-\vec{e}$ is the propagation direction of the incident wave (at the front face of the ellipsoid) and $\vec{e}$ is the propagation of the scattered wave (i.e., the part that will propagate back to the transducer). The vector $\vec{e}$, corresponding to the position $\vec{r}$, is given by

$$\vec{e} = \frac{1}{2} h \nabla\phi = h\overleftrightarrow{Q} \cdot \vec{r} \tag{B-4}$$

from which we deduce

$$\vec{r} = \frac{1}{h} \overleftrightarrow{Q}^{-1} \cdot \vec{e} \ . \tag{B-5}$$

In the above expression h is a normalization factor, as yet undetermined.

The condition that $\vec{r}$ lies on the surface of the ellipsoid is

$$\begin{aligned}1 = \phi(\vec{r}) &\equiv \vec{r} \cdot \overleftrightarrow{Q} \cdot \vec{r} \\ &= \frac{1}{h^2} \vec{e} \cdot \overleftrightarrow{Q}^{-1} \cdot \vec{e}\end{aligned} \tag{B-6}$$

from which we infer

$$h = (\vec{e} \cdot \overleftrightarrow{Q}^{-1} \cdot \vec{e})^{1/2} \ . \tag{B-7}$$

It is obvious that the distance from the center of the ellipse to the front-face tangent plane is

$$d = \vec{r} \cdot \vec{e} = h\vec{r} \cdot \overleftrightarrow{Q} \cdot \vec{r} = h \ . \tag{B-8}$$

Thus we finally obtain the desired result

$$\begin{aligned}d &= d(\vec{e};x) \\ &= (\vec{e} \cdot \overleftrightarrow{Q}^{-1} \cdot \vec{e})^{1/2} \ .\end{aligned} \tag{B-9}$$

In the case of the spheroid we obtain

$$d = [a^2 - (a^2-c^2)(\vec{w}\cdot\vec{e})^2]^{1/2} \ . \tag{B-10}$$

If we write

$$\vec{w} = \vec{e}_x\gamma_x + \vec{e}_y\gamma_y + \vec{e}_z(1-\gamma_x^2-\gamma_y^2)^{1/2} \ , \tag{B-11}$$

then

$$\vec{w}\cdot\vec{e} = \vec{e}\cdot\vec{e}_x\gamma_z + \vec{e}\cdot\vec{e}_y\gamma_y + \vec{e}\cdot\vec{e}_z(1-\gamma_x^2-\gamma_y^2)^{1/2} \ . \tag{B-12}$$

APPENDIX C

DISTRIBUTIONAL PROPERTIES OF ESTIMATORS WITH RESPECT TO MEASUREMENT ERROR WHEN BAYES RULES ARE EMPLOYED

The estimation rule for $\hat{x}$ and $\hat{g}$ is Bayes in that it minimizes the risk with respect to the joint distribution of x, g, and y. As discussed in the third section of the paper, if the prior on x and g is flat, then $\hat{x}_g(y)$ will be the maximum likelihood estimate of the vector parameter x for each postulated inclusion g. In this situation, the distributional properties of $\hat{x}_g(y)$ are considered with respect to the conditional distribution of y given x and g. Even if the prior is not taken as flat, because $\hat{x}_g(y)$ is a function of y, it is still appropriate to consider its distributional properties with respect to this conditional distribution. To do this we note that $\hat{x}_g$ and y are such that the surface $\phi(x,g;y)$ has a stationary point. Defining the matrix P as the Jacobian matrix of $h_g(x)$ with respect to x, that is

$$[P_{ki}] = \left[\frac{\partial h_{g_k}(x)}{\partial x_i}\right]. \tag{C-1}$$

It is easily seen that since

$$0 = \left.\frac{\partial \phi}{\partial x}\right|_{\hat{x}_g,y} = \left.\frac{\partial \phi}{\partial x}\right|_{x_{0_g},y_0} + \left[\frac{\partial^2 \phi}{\partial x_i \partial x_j}\right][\hat{x}_g - x_{0_g}] + \frac{\partial^2 \phi}{\partial y \partial x}(y - y_0) + \cdots , \tag{C-2}$$

then

$$(\hat{x}_g - x_{0_g}) \simeq H_g^{-1} P' C^{-1} (y - y_0) , \tag{C-3}$$

where x_{0_g} is the state that minimizes $\phi(x,g;y_0)$ when $y=y_0$ is observed as measurement error free. From (C-3) it is seen that, since $y_0 = E(y|x_{0_g},g) = h_g(x_{0_g})$, then

$$E(\hat{x}_g|x_{0_g},g) = x_{0_g} \tag{C-4}$$

and

$$\Sigma_{\hat{x}_g|x_{0_g},g} = H_g^{-1} P' C^{-1} P H_g^{-1} . \tag{C-5}$$

We note that the Hessian has the form

$$H_g = \left[\frac{\partial^2 \phi}{\partial x_i \partial x_j}\right] = P' C^{-1} P - Q , \tag{C-6}$$

where

$$[Q_{ij}] = \left[\frac{\partial^2 \ell n\, p(x,g)}{\partial x_i \partial x_j}\right] \tag{C-7}$$

is the Hessian of the log of the prior distribution. In the case that p(x,g) is flat in x, then

$$\Sigma_{\hat{x}_g|x_{0_g}} = (P' C^{-1} P)^{-1} . \tag{C-8}$$

From (C-3) it is seen that since y is multivariate normal with mean $h_g(x_{0_g})$ and variance covariance matrix C, $\hat{x}_g(y)$ will also be multivariate normal with mean (C-4) and variance covariance (C-5). This result is independent of any *a priori* bias that is assumed.

Under normality of the measurement errors, the distributional properties of the residual sum of squares may be derived. The residual sum of squares is defined by

$$S = (y - h_g(\hat{x}_g))' C^{-1} (y - h_g(\hat{x}_g)) . \tag{C-9}$$

In the case where measurement errors in the A_2 experiments are uncorrelated with those in the d experiments, S has the decomposition

$$S = S_1 + S_2$$

with

$$S_1 = (y_1 - h_{g_1}(\hat{x}_g))' C_1^{-1} (y_1 - h_{g_1}(\hat{x}_g)) , \tag{C-10}$$

$$S_2 = (y_2 - h_{g_2}(\hat{x}_g))' C_2^{-1} (y_2 - h_{g_2}(\hat{x}_g)) . \tag{C-11}$$

The statistics S_1 and S_2 are not independent since they both involve $\hat{x}_g$.

We now develop an approximate distribution for S. The cases for S_1 and S_2 are done similarly. Assume we know the matrix C up to a scale factor θ, and the matrix Q up to the scale factor θ^{-1}, that is

$$C = \theta W , \tag{C-12}$$

$$Q = \theta^{-1} U \tag{C-13}$$

with both W and U completely known. This is equivalent to knowing the relative weights of the measurement errors among the different experiments as well as relative to the prior information. Equation (C-3) can now be written

$$\hat{x}_g - x_{0_g} = (P' W^{-1} P - U)^{-1} P' W^{-1} (y - h_g(x_{0_g})) . \tag{C-14}$$

Thus,

$$\begin{aligned} y - h_g(\hat{x}_g) &= y - h_g(x_{0_g}) - (h_g(\hat{x}_g) - h_g(x_{0_g})) \\ &\simeq (y - h_g(x_{0_g})) - P(\hat{x}_g - x_{0_g}) \\ &= (I - P(P' W^{-1} P - U)^{-1} P' W^{-1})(y - h_g(x_{0_g})) . \end{aligned} \tag{C-15}$$

Therefore,

$$S = \frac{1}{\theta} (y - h_g(x_{0_g}))' B (y - h_g(x_{0_g})) , \tag{C-16}$$

with

$$B = (I - W^{-1}P(P'W^{-1}P-U)^{-1}P')W^{-1}$$

$$\cdot (I - P(P'W^{-1}P-U)^{-1}P'W^{-1}) \ . \qquad (C\text{-}17)$$

Thus the scaler $R = \theta S$, is independent of θ and thus is computable without its knowledge.

In the case of a flat prior, $U = 0$. Then

$$B = (I - W^{-1}P(P'W^{-1}P)^{-1}W^{-1} \ . \qquad (C\text{-}18)$$

Equation (C-16) represents S as a quadratic form in the multivariate Gaussian variable y. In the following, we assume that the value θ_0 is used in calculation of S, and that this may in general be different for the true scale factor θ. The first two moments of S are

$$E(S|x,g) = \text{tr}\ (\frac{1}{\theta_0} BC) = \frac{\theta}{\theta_0} \text{tr}\ (BW) \qquad (C\text{-}19)$$

$$\text{var}(S|x,g) = 2\ \text{tr}\left[(\frac{1}{\theta_0} BC)^2\right] = 2(\frac{\theta}{\theta_0})^2\ \text{tr}\ (BW)^2 \qquad (C\text{-}20)$$

where we have written θ_0 for the value of the scale factor assumed in the calculations and θ for the true value of the scale factor. Equations (C-19) and (C-20) are valid even if y is not Gaussian. We note that in the case of a flat prior, BW is idempotent with tr (BW) = rank (BW) = N - m.

Knowledge of (C-19) and (C-20) up to the unknown scale factor θ allows one to develop a chi-square approximation to its distribution. In particular, we have that the random variable

$$\chi_\nu = \frac{\theta_0}{\theta} \frac{\text{tr}\ (BW)}{\text{tr}\ (BW)^2} S \qquad (C\text{-}21)$$

is approximately a chi-square variate with degrees of freedom

$$\nu = [\text{tr}\ (BW)]^2/\text{tr}\ [(BW)^2] \ . \qquad (C\text{-}22)$$

In the case that the prior is flat, the approximation becomes exact. That is,

$$\frac{R}{\theta} = \frac{\theta_0}{\theta} S \sim \chi^2_{N-m} \ . \qquad (C\text{-}23)$$

The above results allow one to construct $100(1-\alpha)\%$ confidence intervals for θ/θ_0. In particular,

$$\frac{S}{\chi^2_{\nu,1-\alpha/2}} \leqslant \frac{\theta}{\theta_0} \leqslant \frac{S}{\chi^2_{\nu,\alpha/2}} \ . \qquad (C\text{-}24)$$

If such a confidence interval does not contain 1, then this is strong evidence that either the assumed measured error scale factor θ_0 is wrong, or that the assumed flaw type is wrong.

In terms of R, (C-24) may be rewritten as a confidence interval for θ as

$$\frac{R}{\chi^2_{\nu,1-\alpha/2}} \leqslant \theta \leqslant \frac{R}{\chi^2_{\nu,\alpha/2}} \ . \qquad (C\text{-}25)$$

Finally, we remark that if an independent estimate of θ is obtained from replicate experiments, then the above distributional properties of R (even when an *a priori* bias is allowed) can be used to test the goodness-of-fit of the scattering model to the data. In particular, if $\hat{\theta}_0$ is an independent estimate of θ based on ν_0 degrees of freedom, then the statistic

$$F = (R/\nu)/\hat{\theta}_0 \qquad (C\text{-}26)$$

will be an F-statistic with ν and ν_0 degrees of freedom. Large values of this statistic indicate that either the scattering model is wrong or the postulated inclusion type is wrong.